UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 10/693,227 | 10/23/2003 | Jerome R. Bellegarda | 4860P3128 | 2262 |

8791 7590 07/13/2007
BLAKELY SOKOLOFF TAYLOR & ZAFMAN
1279 OAKMEAD PARKWAY
SUNNYVALE, CA 94085-4040

| EXAMINER | |
|---|---|
| RIDER, JUSTIN W | |
| **ART UNIT** | **PAPER NUMBER** |
| 2626 | |
| **MAIL DATE** | **DELIVERY MODE** |
| 07/13/2007 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 10/693,227 | BELLEGARDA, JEROME R. | |
| | Examiner | Art Unit | |
| | Justin W. Rider | 2626 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *23 October 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-112* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-15,17-39,41-63,65-87 and 89-112* is/are rejected.

7)☒ Claim(s) *16,40,64 and 88* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *23 October 2003* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *10/03 (2 sheets)*.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

## DETAILED ACTION

1. This action is responsive to communications: Application filed 23 October 2003. Claims 1-112 are pending.

### *Information Disclosure Statement*

2. The information disclosure statement(s) (IDS) submitted on 23 October 2003 is in compliance with the provisions of 37 CFR 1.97. Accordingly, the examiner has considered the information disclosure statement(s).

### *Specification*

3. The disclosure is objected to because of the following informalities: On page 8 of the specification, the United States Patent Application Number is blank.

Appropriate correction is required.

### *Claim Objections*

4. Claims 8, 13-18, 20, 24, 32, 37-42, 44, 48, 56, 61-66, 68, 72, 80, 85-90, 92 and 96 are objected to because of the following informalities: The above claims recite 'the voice table', however there is no antecedent basis for the voice table. A declaration of a voice table is necessary to overcome the objection. Appropriate correction is required.

### *Claim Rejections - 35 USC § 102*

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

6. Claims 97-98, 101-102, 105-106 and 109-110 are rejected under 35 U.S.C. 102(e) as being anticipated by **Coorman et al. (USPN 6,665,641)** referred to as **Coorman** hereinafter.

Claims 97, 101, 105 and 109: **Coorman** discloses a system of determining discontinuities, comprising:

i. gathering time-domain samples from recorded speech segments (col. 20, lines 17-22, *'The database may directly contain digitally sampled waveforms, or it may include pointers to such waveforms,'* [emphasis supplied]);

ii. extracting features that represent the samples (col. 4, lines 23-25, *'The acoustic join cost is based on a quantization of the mel-cepstrum,'*);

iii. determining a discontinuity between the segments (col. 12, 'Cost Functions for Numeric Features', *'Imprecise linguistic or acoustic knowledge, for example, how big a discontinuity in pitch can be perceived,'*), the discontinuity based on a distance between the features (*'For example, the mismatch of pitch between phones with the same accentuation (either both accented, or both unaccented) in the Transition Cost has a symmetric cost function...'*).

*Claims 98, 102, 106, and 110*: **Coorman** discloses a system as per claims 97, 101, 105 and 109 above, wherein the time-domain samples include pitch periods surrounding a boundary of a phoneme (col. 19, lines 1-9).

### *Claim Rejections - 35 USC § 103*

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 1-8, 19-20, 25-32, 43-44, 49-56, 67-68, 73-80, 91-92, 99-100, 103-104, 107-108 and 111-112 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Coorman** in view of **Michael Banbrook, 'Nonlinear Analysis of Speech From a Synthesis Perspective', A thesis submitted for the degree of Doctor of Philosophy at The University of Edinburgh; October 15, 1996 (Specifically Chapter 4),** referred to as **Banbrook** hereinafter.

*Claims 1, 25, 49 and 73*: **Coorman** discloses a method for analyzing speech for use in synthesis, comprising:

i. extracting portions from time-domain speech segments (col. 5, lines 28-30);

ii. creating feature vectors (col. 5, lines 28-30) that represent the portions in a vector space; and

iii. determining a distance between the feature vectors in the vector space (col. 18, lines 16-19).

However, **Coorman** fails to, but **Banbrook** does specifically disclose wherein features include phase information of the portions (p. 37, *'The data is projected onto a phase space defined by the singular vectors of the data, which can then be partitioned into a signal subspace and a noise subspace.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because it introduces a combination of analysis tools (e.g. time delay embedding, singular value decomposition, correlation dimension, local singular value analysis, Lyapunov spectra and short term prediction properties) and looks in detail at Lyapunov exponents and two major novel modifications are proposed that are demonstrated to be more robust than conventional techniques (Abstract).

Claims 2, 26, 50 and 74: **Coorman** discloses a system as per claims 1, 25, 49 and 73 above, wherein creating feature vectors comprises constructing a matrix W from the portions (col. 18, lines 21-23, *'The calculation of this spectral mismatch is based on a distance calculation between spectral vectors. This might be a heavy task as there can be many segment combinations possible. In order to reduce the computational complexity a combination matrix--containing the spectral distances- could be calculated in advance.'* [emphasis added]).

However, **Coorman** recites performing operations on said matrix, however failing to, but **Banbrook** does specifically disclose decomposing the matrix W (p. 37, *'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

<u>*Claims 3, 27, 51 and 75*</u>: **Coorman** discloses a system as per claims 2, 26, 50 and 74 above, further comprising extracting global boundary-centric features from the portions (col. 10, lines 49-54).

<u>*Claims 4, 28, 52 and 76*</u>: **Coorman** discloses a system as per claims 2, 26, 50 and 74 above, wherein the speech segments each include a segment boundary within a phoneme (col. 9, lines 5-8).

<u>*Claims 5, 29, 53 and 77*</u>: **Coorman** discloses a system as per claims 4, 28, 52 and 76 above, wherein the speech segments each include at least one diphone (col. 9, lines 5-8).

<u>*Claims 6, 30, 54 and 78*</u>: **Coorman** discloses a system as per claims 5, 29, 53 and 77 above, wherein the portions include at least one pitch period (col. 19, lines 7-9).

<u>*Claims 7, 31, 55 and 79*</u>: **Coorman**, in view of **Banbrook** disclose a system as per claims 6, 30, 54 and 78 above. However, **Coorman** fails to, but **Banbrook** does specifically disclose wherein decomposing the matrix *W* comprises performing a pitch synchronous (p. 37, *'which can then be partitioned into a signal subspace and a noise subspace.'*) singular value analysis on the pitch periods of the time-domain segments (p. 37, *'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

*Claims 8, 32, 56 and 80*: **Coorman**, in view of **Banbrook** disclose a system as per claims 6, 30, 54 and 78 above. However, **Coorman** fails to, but **Banbrook** does specifically disclose wherein the matrix $W$ is a $2KM \times N$ matrix represented by

$$W = U\Sigma V^T$$

where $K$ is the number of pitch periods near the segment boundary extracted from each segment, $N$ is the maximum number of samples among the pitch periods, $M$ is the number of segments in the voice table having a segment boundary within the phoneme, $U$ is the $2KM \times R$ (p. 37, N x $w$ trajectory matrix found utilizing time delay embedding) left singular matrix with row vectors $u_i$ $(1 \leq i \leq 2KM)$, $\Sigma$ is the $R \times R$ diagonal matrix of singular values $s_1 \geq s_2 \geq \ldots \geq s_R > 0$, $V$ is the $N \times R$ right singular matrix with row vectors $v_j$ $(1 \leq j \leq N)$, $R << 2KM$, and $^T$ denotes matrix transposition, wherein decomposing the matrix $W$ comprises performing a singular value decomposition of $W$ (p. 37-38, ***X**= **S**ΣC^T, where **X** is the trajectory matrix, **S** and C are the matrices of the singular vectors associated with Σ, which is a diagonal matrix.* ).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

*Claims 19, 43, 67 and 91*: **Coorman** discloses a system as per claims 5, 29, 53 and 77 above, wherein the portions include centered pitch periods (col. 19, lines 7-9, *'In the preferred*

*embodiment the length of the trailing and leading regions are of the order of one to two pitch periods and the sliding window is bell-shaped [i.e. centered].'*).

*Claims 20, 44, 68 and 92*: Claims 20, 44, 68 and 92 are similar in scope and content to that of claim 8 above and so therefore are rejected under the same rationale.

*Claims 99, 103, 107 and 111*: **Coorman** discloses a system as per claims 98, 102, 106, and 110 above. However, **Coorman** fails to, but **Banbrook** does specifically disclose wherein features include phase information of the portions (p. 37, *'The data is projected onto a phase space defined by the singular vectors of the data, which can then be partitioned into a signal subspace and a noise subspace.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

*Claims 100, 104, 108 and 112*: **Coorman** discloses a system as per claims 99, 103, 107 and 111 above, wherein creating feature vectors comprises constructing a matrix W from the portions (col. 18, lines 21-23, *'The calculation of this spectral mismatch is based on a distance calculation between spectral vectors. This might be a heavy task as there can be many segment combinations possible. In order to reduce the computational complexity a combination matrix--containing the spectral distances- could be calculated in advance.'* [emphasis added]).

However, **Coorman** recites performing operations on said matrix, however failing to, but **Banbrook** does specifically disclose decomposing the matrix W (p. 37, *'The method of singular value decomposition (SVD) reduction, described by Broomhead and King [85, 103], addresses this problem.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

Claims 17-18, 41-42, 65-66 and 89-90: **Coorman** discloses a system as per claims 2, 26, 50 and 74 above, wherein said distances are associated with said speech segments (units, col. 11, section 'Cost Functions', lines 46-49, *'a set of nonlinear cost functions has been defined for use in the unit selection...with specific properties which help in the unit selection process.'*).

9. Claims 9-10, 21-23, 33-34, 45-47, 57-58, 69-71, 81-82 and 93-95 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Coorman**, in view of **Banbrook** and in further view of **Ansari et al., 'Pitch Modification of Speech Using a Low-Sensitivity Inverse Filter Approach'; IEEE Signal Processing Letters; March 1998** referred to as **Ansari** hereinafter.

Claims 9, 33, 57 and 81: **Coorman**, in view of **Banbrook** disclose a system as per claims 8, 32, 56 and 80 above, however failing to, but **Ansari** does specifically disclose padding a signal with zeroes (p. 61, section III, *'in the new method when the residual is modified with zero-padding to lower the pitch.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Ansari** in the system of **Coorman**, in view of **Banbrook** because speech modifications using the method of **Ansari** are superior in quality to those obtained with RELP, while at the same time being less sensitive than RELP to errors in pitch marking (Abstract).

*Claims 10, 34, 58 and 82*: **Coorman**, in view of **Banbrook** disclose a system as per claims 9, 33, 57 and 81 above. However, **Coorman** fails to, but **Banbrook** does specifically disclose wherein a feature vector $u_i$ is calculated as

$$\bar{u}_i = u_i \Sigma$$

where $u_i$ is a row vector associated with a pitch period $i$, and $\Sigma$ is the singular diagonal matrix (p.49, *'In general, any matrix A can be written A = QR (4.17) where Q has orthogonal columns and R is a square upper-right triangular matrix with positive values on the diagonal.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Banbrook** in the system of **Coorman** because of the reasons described above.

*Claims 21, 45, 69 and 93*: **Coorman**, in view of **Banbrook** disclose a system as per claims 20, 44, 68 and 92 above, however failing to, but **Ansari** does specifically disclose symmetrically padding a signal with zeroes (p. 61, section III, *'in the new method when the residual is modified with zero-padding to lower the pitch.'*). It would have been obvious to one having ordinary skill in the art that if pitch periods were centered, that one would be motivated to append zeros symmetrically on either side of the centered samples in order to maintain symmetric proportions with respect to a centered pitch.

*Claims 22-23, 46-47, 70-71 and 94-95*: Claims 22-23, 46-47, 70-71 and 94-95 are similar in scope and content to that of claims 10 and 12 above and so therefore are rejected under the same rationale.

10. Claims 11-15, 35-39, 59-63 and 83-86 are rejected under 35 U.S.C. 103(a) as being unpatentable over **Coorman** and **Banbrook** in view of **Ansari** and in further view of **Jerome R. Bellegarda, 'Exploiting Latent Information in Statistical Language Modeling'** referred to as **Bellegarda** hereinafter.

*Claims 11, 35, 59 and 83*: **Coorman** and **Banbrook** in view of **Ansari** disclose a system as per claims 10, 34, 58 and 82 above. However, **Coorman** and **Banbrook** in view of **Ansari** fail to, but **Bellegarda** does specifically disclose wherein the distance between two feature vectors is determined by a metric comprising the cosine of the angle between the two feature vectors (p. 5, *'We conclude that a natural metric to consider for the "closeness" between words is therefore the cosine of the angle between $\bar{u}_i$ and $\bar{u}_j$.'*).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Bellegarda** in the system of **Coorman** and **Banbrook** in view of **Ansari** because it uses latent semantic analysis to improve statistical language modeling utilizing existing clustering techniques capable of aiding in speech production by machine (Introduction).

*Claims 12, 36, 60 and 84*: **Coorman** and **Banbrook** in view of **Ansari** disclose a system as per claims 10, 34, 58 and 82 above. However, **Coorman** and **Banbrook** in view of **Ansari** fail to, but **Bellegarda** does specifically disclose wherein the metric comprises a closeness measure, $C$, between two feature vectors, $\bar{u}_k$ and $\bar{u}_l$, wherein $C$ is calculated as

$$C(\bar{u}_k, \bar{u}_l) = \cos(u_k\Sigma, u_l\Sigma) = \frac{u_k \Sigma^2 u_l^T}{\| u_k \Sigma \| \; \| u_l \Sigma \|}$$

for any $1 \leq k, l \leq 2KM$ (p. 6, (10)).

Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include the teachings of **Bellegarda** in the system of **Coorman** and **Banbrook** in view of **Ansari** because of the reasons described above.

*Claims 13, 37, 61 and 85*: **Coorman** and **Banbrook** in view of **Ansari** and in further view of **Bellegarda** disclose a system as per claims 12, 36, 60 and 84 above. The examiner is taking Official Notice that the difference as calculated by

$$d(S_1,S_2) = d_0(p_1, q_1) = 1 - C(\bar{u}_{p1}, \bar{u}_{q1})$$

is simply a natural extension from the closeness measure as determined in the prior claim (which is assumed to be a value between 0 and 1). Therefore, it would have been obvious to one having ordinary skill in the art at the time of invention to include a difference measure as the counterpart to a closeness measure previously determined because it is well known that the two factors have an inversely variable relationship, here adding up to 1.

*Claims 14, 38, 62 and 86*: **Coorman** discloses a system as per claims 13, 37, 61 and 85 above, wherein the calculation for the difference between two segments in the voice table, S1 and S2, is expanded to include a plurality of pitch periods from each segment (col. 19, lines 1-9).

*Claims 15, 39, 63 and 87*: **Coorman** discloses a system as per claims 13, 37, 61 and 85 above, wherein the difference between two segments in the voice table, $S_1$ and $S_2$, is associated with a discontinuity between $S_1$ and $S_2$ (col. 18, lines 48-54, *'The major concern of waveform concatenation is in avoiding waveform irregularities such as discontinuities and fast transients that may occur in the neighborhood of the join...It is thus important to minimize signal discontinuities at each junction.'*).

***Allowable Subject Matter***

11. Claims 16, 40, 64 and 88 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The above claims recite a specific distance formula as follows:

$$d(S_1,S_2) = \left| d_0(p_1,q_1) - \frac{d_0(p_1,\bar{p}_1) + d_0(q_1,\bar{q}_1)}{2} \right| = \left| \frac{C(\bar{u}_{p1},\bar{u}_{\bar{p}1}) + C(\bar{u}_{q1},\bar{u}_{\bar{q}1})}{2} - C(\bar{u}_{p1},\bar{u}_{q1}) \right|$$

The above is used as an alternative distance measure that is essentially the relative change in similarity that occurs during a concatenation function. More specifically, this alternative distance specifically shows wherein a difference is zero only when two identical segments are concatenated together; otherwise a difference measure greater than zero exists. While the cited prior art references do use distance measures as disclosed, none of the references use an alternative distance measure as specifically disclosed and defined as per claims 16, 40, 64 and 88.

***Conclusion***

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. **Marcus (USPN 4,813,074), Abe et al. (USPN 5,581,652), Braida et al. (USPN 5,745,843) and Beyerlein et al. (USPN 5,933,806)** disclose clustering techniques utilizing feature vector distances; **Hermansky et al. (USPN 5,537,647), Holzapfel (US 2002/0035469 A1), Tzirkel-Hancock (USPN 6,275,795)** perform speech signal segmentation for various applications.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Justin W. Rider whose telephone number is (571) 270-1068. The examiner can normally be reached on Monday - Friday 7:30AM - 5:00PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600

J.W.R.
12 June 2007